School Redistricting: Wiping Unfairness Off the Map

Ariel D. Procaccia†, Isaac Robinson‡, and Jamie Tucker-Foltz†

†Harvard University
‡University of Oxford

Abstract

We introduce and study the problem of designing an equitable school redistricting map, which we formalize as that of assigning n students to school attendance zones in a way that is fair to various demographic groups. Drawing on methodology from fair division, we consider the demographic groups as players and seats in schools as homogeneous goods. Due to geographic constraints, not every school can be assigned to every student. This raises new obstacles, rendering some classic fairness criteria infeasible. Nevertheless, we show that it is always possible to find an almost proportional allocation among g demographic groups if we are allowed to add $O(g \log g)$ extra seats. For any fixed g, we show that such an allocation can be found in polynomial time, obtaining a runtime of $O(n^2 \log n)$ in the special (but practical) case where $g \leq 3$.

1 Introduction

In the United States and around the world, the public school a student attends is typically determined by their home address; the area associated with a particular school is called a *school attendance zone*.[1] Since some public schools are considered significantly better than others even within the same municipality, the boundaries of school attendance zones have an outsize impact on the lives of students. This is reflected, for example, in the willingness of parents to pay a premium for living in a good school attendance zone [Black, 1999].

Every now and then, municipalities have to redraw their school attendance zones — a process known as *school redistricting* — due to demographic shifts, changes in the capacity of schools (the most extreme of which occur when schools open or close), or other reasons. This process is almost always controversial and often harrowing. As an example, in 2019, Howard County in Maryland adopted a school redistricting plan that aimed to "balance school capacity utilization, provide relief to schools most impacted by crowding, and address inequities in the distribution of students affected by poverty."[2] According to the New York Times,[3]

> "[That redistricting plan] has led to bitter divisions. Protesters in matching T-shirts have thronged school board meetings. Thousands of letters and emails opposing the

[1] In some cities in the US, such as New York City and Boston, there is a city-wide *school choice* process matching students to (public) schools. Although there is a large literature on school choice [Abdulkadiroğlu et al., 2005a,b], it is practiced in only a small number of cities.

[2] https://www.hcpss.org/school-planning/boundary-review

[3] https://www.nytimes.com/2019/11/12/us/howard-county-school-redistricting.html

> redistricting plan, some of them overtly racist, have poured in to policymakers. One high school student made a death threat against the superintendent of schools, Michael J. Martirano."

In the hope of (eventually) alleviating the strife around school redistricting, we draw on the field of *fair division* and seek to generate redistricting plans that are *provably fair*. As is often the case in fair division, the challenge is actually twofold: We must *identify feasible notions of fairness for school redistricting, and design (computationally tractable) algorithms for achieving them.* To our knowledge, this is the first paper on fair division for school redistricting, and as such it does not presume to entirely capture the real-world nuance of the problem; our goal is to gain an initial understanding of the possibility of provable fairness in this domain.

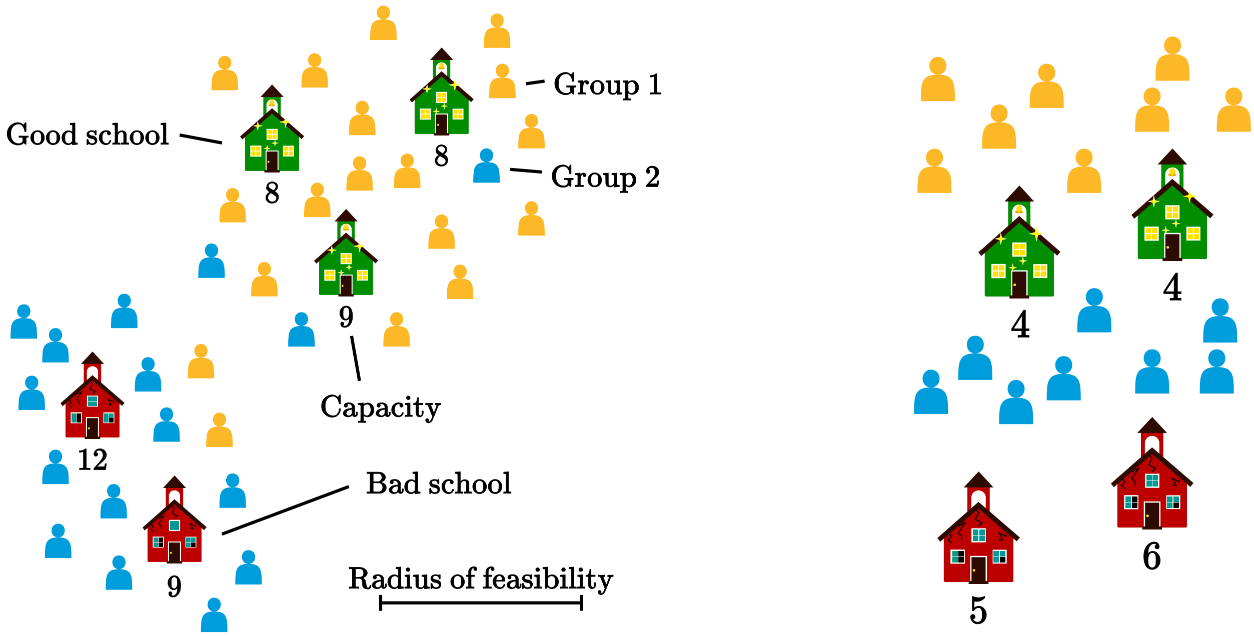
1.1 Our Approach

We consider a set of students and a set of schools, each with a given capacity. A key assumption we make is that there is an agreed metric by which the quality of schools is measured, given, say, by their scores according to a website like *Niche* or *GreatSchools*. Granted, the methodology for computing such scores is notoriously subjective and different families will have different preferences. However, this metric is not used to determine the actual assignment of students — it is merely a tool for measuring the fairness of school attendance zones in terms of aggregate school quality, as we explain later, and as such we believe it is practical.

In addition, there are feasibility constraints that restrict the set of schools a given student can be assigned to; in practice, these would typically be based on the distance between the student and the school. While there are a few municipalities that do not impose such restrictions, it is common practice to design school attendance zones in a way that limits distances; this is advantageous both in terms of the cost of transportation services and the educational experience of students, as long bus rides have been tied to reduced attendance and chronic absenteeism [Cordes et al., 2022].

Concerns about equity in the context of school redistricting usually arise when racial groups are seen as disadvantaged by the quality of schools available to them. Let us, therefore, consider a partition of the students into demographic groups; these demographic groups are the "players" in our fair division problem. Given such a group and a subset of seats in schools, imagine a bipartite graph between students and seats, where there is an edge between a student and a seat if the corresponding school is feasible for the student, and the weight on an edge represents the quality of the school; the utility of the group for the subset of seats is given by the maximum weight matching in this graph. This means that the utility of a group of students is additive, summing up the quality of the schools each student attends.

In order to gain some intuition for how we might define fairness to demographic groups, let us first identify a few examples of unfairness that will *not* be addressed by our framework, depicted in Figure 1. First, imagine that a group is concentrated in an area that has only "bad" schools, with no "good" schools in the vicinity, as is the case for the blue group in Figure 1a. To satisfy the feasibility constraints, any assignment would have to give Group 2 low utility. A possible solution may be to build new schools or invest more in the current ones, but these remedies are out of scope for school redistricting. Second, consider the scenario in Figure 1b. Here, the two demographic groups are situated in the north and the south of the city; there is a cluster of "good" schools in the center that are feasible for both groups and a cluster of "bad" schools in the far south that are only feasible for Group 2. Furthermore, the number of seats in the central schools is equal to the



(a) Group 2 is concentrated around "bad" schools.

(b) Matching group 2 to good schools would block group 1.

Figure 1: Two inherently unfair situations where group 2 students must always be allocated mostly to "bad" schools due to distance and capacity constraints.

number of students in Group 1. While students in Group 2 can feasibly be assigned to seats in the central schools, that would mean that some students in the north group cannot go to school at all, which is a nonstarter. To lodge a justified complaint against a school redistricting plan, therefore, a group has to demonstrate a alternative plan where *all students are assigned* and its members are better off.

With this insight in mind, we can now adapt standard notions of fairness. Specifically, we focus on two central notions in fair division, which are both prevalent, for example, in the literature on cake cutting [Procaccia, 2016]. The first notion, *proportionality*, typically means that each of the n players receives at least $1/n$ of their utility for the entire set of goods (or the whole cake). In our setting, we say that an allocation of students to schools with g groups is *proportional* if each group receives at least $1/g$ of its maximum utility under any *feasible* allocation. The second notion, *envy-freeness*, typically says that no player wants to swap their allocation with another player (so no player envies another). In the context of school redistricting, we say that an allocation is *envy-free* if, for any two groups, there is no alternative *feasible* allocation where one group receives a subset of the seats the other group was allocated and achieves higher utility.

We consider additional relaxations of these notions. First, in the style of the algorithmic literature on *resource augmentation* [Roughgarden, 2020], we allow the capacity constraints to be violated, adding a number of extra seats to some schools. A minor capacity violation is practical; it can be handled, for example, by setting aside a small budget for handling overflow, or (if one is conservative) by computing a solution for capacities that are slightly lower than their true values. Second, inequalities over utilities are required to hold only up to an additive error. This is inspired by existing relaxed notions, namely proportionality up to one good [Conitzer et al., 2017] and envy-freeness up to one good [Budish, 2011, Caragiannis et al., 2019].

1.2 Our Results and Techniques

In Section 3, we show that to achieve even a constant-factor approximation to envy-freeness, capacity constraints must be increased by a constant fraction of the *total* capacity. We take this result to mean that envy-freeness is a nonstarter in our setting.

By contrast, our results for proportionality are very encouraging. Specifically, we say that an allocation is *t-proportional* if it violates capacities by at most t seats, and the utility of each of the g groups is at least $1/g$ of its maximum utility (as described above), up to the value of t seats. We wish to bound t as a function of the number of groups g, not the number of students, schools, or the structure of the underlying feasibility graph.

It is not too difficult to construct instances in which $t = \Omega(g)$. The first main result of this paper is an algorithm that computes t-proportional allocations for $t = O(g \log g)$. The proof, which we present in Section 4, is via a recursive reduction of the problem to the continuous fair division problem of *consensus halving* [Simmons and Su, 2003]. In broad strokes, we divide the groups into two almost-equal sets and recursively find an almost-proportional school assignment for each set individually, ignoring the utilities of the groups in the other set. This yields two school redistricting plans, each with a small number of school capacity violations. We view these plans as matchings on a bipartite graph between students and school seats, and consider the subgraph C that is the union of the two matchings. Since the maximum degree of any vertex is two, C decomposes into a disjoint union of paths and cycles, which we arrange in a large cycle to form a circular "cake." We then apply a consensus halving protocol on this cake to find a compromised plan between the two matchings, where each "cut" of the cake corresponds to an additional capacity violation, and perhaps an additional small loss in utility.

While we view the main contribution of this paper as establishing the existence of almost-proportional allocations in a novel setting, our proof also shows that such allocations can be found in polynomial time for any constant number of groups g. However, the time complexity that comes out of the proof remains in practice large: $O(n^{4g-2})$ for all $g \geq 2$. In applied settings, we expect g to be at least 2 or 3, which would lead to respective time complexities of $O(n^6)$ and $O(n^{10})$. Already, this could be well out of the range of feasibility for school redistricting instances on a real-world scale. Our second main result, presented in Section 5, is an alternative proof for the cases where $g \in \{2, 3\}$ yielding a smaller total number of school capacity violations, a better utility guarantee, and, most importantly, an improved runtime of $O(n^2 \log n)$ for both cases. For the case of $g = 3$, this involves reducing our problem to a novel variant of the *mountain climbing problem* [Goodman et al., 1989], in which two climbers must summit a 2-dimensional mountain from opposite directions while always maintaining the same elevation. For well-behaved mountains, a route always exists, which we use to show the existence of an almost-proportional allocation that can be computed efficiently.

1.3 Related Work

Prior work on school redistricting centers on optimization (especially mathematical programming) approaches [Clarke and Jurkis, 1968, Liggett, 1973, Franklin and Koenigsberg, 1973, Caro et al., 2004, Allman et al., 2022, Gillani et al., 2023]. Most relevant is the work of Allman et al. [2022], who report on a collaboration with the San Francisco Unified School District on its redistricting plan.

Their integer program minimizes the maximum shortage of seats in any zone subject to roughly equal zone sizes, contiguity of zones, school capacity constraints, and student diversity constraints. Their work serves as an encouraging proof of concept that municipalities can embrace algorithmic approaches for school redistricting.

The problem of *political redistricting* — dividing a state into Congressional or state legislative districts — is somewhat more rarefied than school redistricting, yet it has received significantly more attention. In particular, there are several papers taking a fair division approach to this problem [Landau et al., 2009, Landau and Su, 2014, Pegden et al., 2017, De Silva et al., 2018, Brams, 2020, Benadè et al., 2023]. The problem is quite different, however, as fairness is defined with respect to parties and the number of seats they would win (i.e., the number of districts where they would have a majority) under various maps. For example, the *geometric target* criterion [Landau and Su, 2014] requires that each party win at least a number of seats equal to the average of its seats under the best and worst possible maps (rounded down); Benadè et al. [2023] show that this criterion is feasible in a cake-cutting-inspired model. Agarwal et al. [2022] and Ko et al. [2022] consider a different notion of fairness for political redistricting based on a popular solution concept known as the core; it is based on eliminating potential deviations of groups that are large enough to form (contiguous) districts.

In fair division, in addition to results used directly and discussed above and below, the most closely related work [Benabbou et al., 2019, 2020] deals with so-called OXS valuations, a subset of submodular valuations. Under OXS valuations, there is a weighted bipartite graph with individuals on one side and items on the other, where the weights correspond to values; the utility of a group of individuals for a subset of items is the weight of the maximum weight matching between them. Our problem is easier in some ways and harder in others: it is easier because all students have identical values for any particular school, but it is harder because all students must be matched and the possible assignments of students to schools are constrained. While fair division problems under various kinds of constraints have been studied [Suksompong, 2021], existing results are inapplicable to our setting.

2 Model and Fairness Notions

An instance of the *school redistricting problem* is specified by:

- A set of students $N = \{1, 2, \ldots, n\}$.
- A partition of the students into g groups, $N = N_1 \cup N_2 \cup \cdots \cup N_g$.
- A set of schools $M = \{1, 2, \ldots, m\}$, where each school k has capacity c_k and value v_k.
- A bipartite graph $\mathcal{G} = (N \cup M, E)$, where the edge relation E describes which schools are feasible for which students. For the purposes of runtime analysis, we assume the student side of the graph has bounded average degree and all schools are feasible for at least one student.

Allocations. An *allocation* is a function $A : N \to M$ such that, for any student $j \in N$, $\{j, A(j)\} \in E$. The *utility* of group i with respect to allocation A is

$$u(i, A) := \sum_{j \in N_i} v_{A(j)}.$$

The *capacity violation* of A is

$$cv(A) := \sum_{k \in M} \max\{0, |\{j \in S \mid A(j) = k\}| - c_k\}.$$

We say A *satisfies capacities* if $cv(A) = 0$. An instance of the school redistricting problem is *feasible* if there is an allocation that satisfies capacities.

Fairness notions. Even when there are no feasibility constraints, one cannot always hope to get a proportional or envy-free allocation due to the fact that seats are indivisible. Thus, we have to relax these notions, considering proportionality/envy-freeness up to an additive error. In our constrained setting, we find that we additionally have to relax the capacity constraints slightly. Our fairness definitions bundle both relaxations together.

For any nonnegative integer t, we say an allocation A is *t-proportional* if $cv(A) \leq t$ and, for any group $i \in G$ and any alternative allocation A' that satisfies capacities,

$$u(i, A) \geq \frac{u(i, A')}{g} - t \max_{1 \leq k \leq m} v_k.$$

We say A is *t-envy-free* if $cv(A) \leq t$ and, for all groups $i_1, i_2 \in G$, there is no alternative allocation A' such that A' satisfies capacities,

$$u(i, A') > u(i, A) + t \max_{1 \leq k \leq m} v_k,$$

and for any school $k \in M$,

$$\left|\{j \in N_{i_1} \mid A'(j) = k\}\right| \leq |\{j \in N_{i_2} \mid A(j) = k\}|.$$

In words, this final condition states that the seats allocated to the envious party i_1 in the new allocation can be viewed as a subset of what was allocated to envied party i_2 in the original allocation. Note that this is a very stringent requirement, placing a high burden of proof on i_1 to be able to claim that i_2 is getting preferential treatment; thus we might expect it to be easy to find an envy-free allocation. However, the only result about envy-freeness in this paper is negative, meaning having a weak notion of envy-freeness only makes its infeasibility stronger.

3 Lower Bounds

In this section, we establish lower bounds for proportionality and envy-freeness in our model. Since the proofs are simple, these results may also help build intuition for the model and definitions.

We start with a lower bound for t-envy-freeness. As mentioned previously, since it requires t to be a constant fraction of the total capacity, we view it as a showstopper.

Theorem 1. *For any n' and C', there is a feasible instance of the school redistricting problem with three groups, $n \geq n'$ students, and total capacity $C \geq C'$, in which there is no t-envy-free allocation for any $t < C/9$.*

Proof. Given n' and C', let n be a number divisible by 3 such that $n \geq n', C'$, and let $C = n$. Consider three groups, each of size $n/3$, and three schools, each with capacity $n/3$. Two of the schools have value 1 ("good" schools), whereas the third has value 0 ("bad" school). The students in group 1 can be assigned only to the two "good" schools, whereas the students in groups 2 and 3 can be assigned to any school. See Figure 2a for an illustration of the case of $n = 12$.

Suppose that A is an allocation such that $cv(A) < \frac{C}{9}$. We will show that some group is envious by more than $\frac{C}{9}$. It must be the case that group 1 is entirely assigned to the "good" schools due to the feasibility constraints. Therefore, since $cv(A) < C/9$, the number of seats in "good" schools available to groups 2 and 3 is smaller than

$$\frac{2n}{3} - \frac{n}{3} + \frac{n}{9} = \frac{4n}{9}.$$

It follows that at least one of these groups — without loss of generality, group 2 — has fewer than $2n/9$ students assigned to the "good" schools, so $u(2, A) < 2n/9$.

We now argue that group 2 envies group 1 by more than $C/9 = n/9$. Indeed, consider an alternative allocation A' where group 2 receives the assignment of group 1 under A; formally, there is a bijection π from N_2 to N_1 such that for all $j \in N_2$, $A'(j) = A(\pi(j))$. In addition, A' assigns group 1 to the remaining "good" schools and group 3 entirely to the "bad" school. We have that

$$u(2, A') = \frac{n}{3} = \frac{2n}{9} + \frac{n}{9} > u(2, A) + \frac{n}{9},$$

and group 2 receives a (weak) subset of the assignment of group 1 under A. Thus, group 2 is envious by more than $\frac{n}{9}$. □

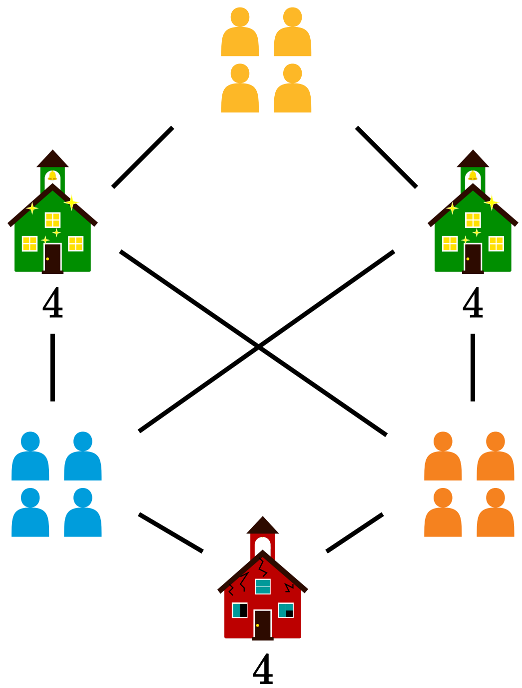
By contrast, we view the following lower bound for proportionality as a minor inconvenience because the capacity violation and utility gap are a small additive term. This opens the door to the positive results that form the backbone of the paper.

Theorem 2. *For any g, n', and C', there is a feasible instance of the school redistricting problem with g groups, $n \geq n'$ students, and total capacity $C \geq C'$ in which there is no t-proportional allocation for any $t < \lfloor \frac{g}{2} \rfloor$.*

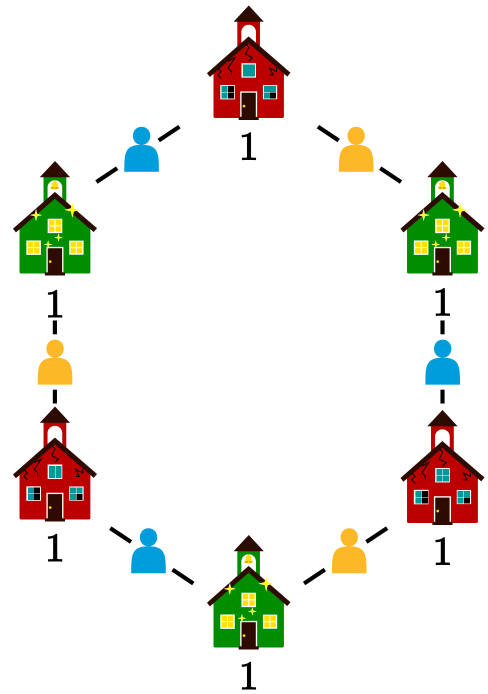
Proof. Consider first the case of $g = 2$ groups. Given n' and C', let n be a number divisible by 2 such that $n \geq n', C'$, and let $C = n$. Each group includes $n/2$ students, and there are $n/2$ schools with value 1 ("good" schools) and $n/2$ schools with value 0 ("bad" schools); all schools have capacity 1. The feasibility graph $\mathcal{G}$ arranges the students and schools in a cycle with repeating segments, each consisting of a "bad" school, student from group 1, "good" school, and student from group 2. See Figure 2b for an illustration of the case of $n = 6$.

For each group i, there is an allocation A_i that matches the students in N_i only to "good" schools, that is, $u(i, A_i) = n/2$. Therefore, to be t-proportional, an allocation A must satisfy $u(i, A) \geq n/4 - t$ for $i = 1, 2$. Note that this is nonzero for sufficiently large n and fixed $t < \lfloor \frac{g}{2} \rfloor$. However, any allocation that does not violate capacities must either assign all students in group 1 to "good" schools and all students in group 2 to "bad" schools, or vice versa, so any such allocation A would have $u(i, A) = 0$ for one group i. We conclude that violating the capacities by at least 1 is necessary to achieve proportionality for two groups.

To extend this construction to the general case of g groups, we create $\lfloor g/2 \rfloor$ cycles as above, each with students from two "new" groups. By the same reasoning as before, proportionality requires a capacity violation of 1 per cycle, and the overall capacity violation must be at least $\lfloor g/2 \rfloor$. □



(a) Construction from the proof of Theorem 1 for $n = 12$.



(b) Construction from the proof of Theorem 2 for $n = 6$.

Figure 2: Counterexamples to the existence of envy-free and proportional allocations.

4 Existence of Almost-Proportional Allocations

In this section we prove the following theorem.

Theorem 3. *In any instance of the school redistricting problem with g groups, there exists an $O(g \log g)$-proportional allocation that can be found in polynomial time for any constant number of groups g.*

A key step in the proof is to appeal to an existence result from continuous fair division, namely, for the problem of *consensus halving*. In this problem, a heterogeneous group of agents must divide a cake into two pieces that they all agree have equal value. We begin by introducing the following novel variant of this problem that is closer to our application.

An instance of the *cake-frosting problem* is a tuple $(f_1, f_2, \ldots, f_g, r)$, where each function $f_i : [0, 1] \to \mathbb{R}$ and $0 < r < g$. We assume the instance satisfies the following axioms.

(1) Each f_i is integrable and non-atomic. If X is a finite union of intervals (meaning connected subsets of $[0, 1]$), we write

$$v_i(X) := \int_{x \in X} f_i(x) dx.$$

(2) Each f_i is piecewise constant (note that this implies (1); we list it separately for the purposes of exposition).

(3) For all $x \in [0,1]$,

$$\sum_{i=1}^{g} f_i(x) = 0.$$

A *perfect frosting* is a finite union of intervals $X \subseteq [0,1]$ such that, for each $1 \leq i \leq g$,

$$v_i(X) = \frac{r}{g} v_i([0,1]).$$

The culinary analogy is that we have n players with competing preferences over which parts of the cake should and should not be frosted. We wish to frost part of the cake to simultaneously interpolate, for each player i, an r-fraction of the way from their utility under a completely unfrosted cake to their utility under a completely frosted cake. We define the number of *cuts* of a perfect frosting X to be the number of boundary points between frosted and unfrosted cake, i.e., twice the minimum number of disjoint intervals needed to express X, minus one for each element of $\{0,1\}$ that is in X.

The cake-frosting problem differs from consensus halving in the following ways:

- [Harder] The density functions may take both positive and negative values.
- [Harder] The density functions need not integrate to 1.
- [Harder] The prescribed values of the two pieces are not necessarily equal.
- [Easier] We have additional axioms (2) and (3).

In the proof of Theorem 3, the frosted part of the cake will correspond to one allocation and the unfrosted part to a different allocation. At the boundaries between these two pieces some schools may be over-enrolled. Hence, in order to minimize the overall capacity violation, we seek to minimize the number of such boundaries, or "cuts" as they are often referred to in the cake-cutting literature. The following lemma uses a classic result in fair division, the Stromquist-Woodall Theorem [Stromquist and Woodall, 1985], to show that perfect frostings exist with few cuts.

Lemma 4. *Any g-player instance of the cake-frosting problem admits a perfect frosting with at most $4g-2$ cuts.*

Proof. Let $(f_1, f_2, \ldots, f_g, r)$ be an instance of the Cake-Frosting problem. For each $1 \leq i \leq g$, let

$$f_i^1(x) := \max\{f_i(x), 0\}, \qquad f_i^2(x) := -\min\{f_i(x), 0\},$$

and, for each $j \in \{1,2\}$, let

$$T_i^j := \int_0^1 f_i^j(x)dx.$$

Let

$$S := \{(i,j) \mid 1 \leq i \leq g,\ j \in \{1,2\},\ T_i^j > 0\}.$$

Then, for each $(i,j) \in S$, the function

$$\mu_i^j(x) := \frac{f_i^j(x)}{T_i^j}$$

is a measure on $[0,1]$. By Theorem 1 of Stromquist and Woodall [1985], there exists a set X that can be written as a union of at most

$$|S| - 1 \leq 2g - 1$$

intervals (which is at most $4g-2$ cuts) such that, for each $(i,j) \in S$,

$$\int_{x \in X} \mu_i^j(x)dx = \frac{r}{g} \int_0^1 \mu_i^j(x)dx.$$

Then, for each i,

$$\begin{aligned}
v_i(X) &= \int_{x \in X} f_i(x)dx \\
&= \sum_{j \in \{1,2\}} \int_{x \in X} f_i^j(x)dx \\
&= \sum_{\substack{j \in \{1,2\} \\ (i,j) \in S}} T_i^j \int_{x \in X} \mu_i^j(x)dx \\
&= \sum_{\substack{j \in \{1,2\} \\ (i,j) \in S}} \frac{T_i^j r}{g} \int_0^1 \mu_i^j(x)dx \\
&= \sum_{\substack{j \in \{1,2\} \\ (i,j) \in S}} \frac{T_i^j r}{g} \cdot \frac{1}{T_i^j} \int_0^1 f_i^j(x)dx \\
&= \frac{r}{g} \sum_{\substack{j \in \{1,2\} \\ (i,j) \in S}} \int_0^1 f_i^j(x)dx \\
&= \frac{r}{g} \int_0^1 f_i(x)dx \\
&= \frac{r}{g} v_i([0,1]),
\end{aligned}$$

so X is a perfect frosting. □

Lemma 5. *Suppose there is a function $q : \mathbb{Z}_{\geq 1} \to \mathbb{Z}_{\geq 1}$ such that any g-player instance of the cake-frosting problem admits a perfect frosting with $q(g)$ cuts. Then, for any g-group instance of the school redistricting problem, there exists an $h(g)$-proportional allocation, where $h(g)$ is defined recursively by*

$$\begin{aligned}
&h(1) = 0, \\
&h(g) = h\left(\left\lceil \frac{g}{2} \right\rceil\right) + h\left(\left\lfloor \frac{g}{2} \right\rfloor\right) + 4q(g) \qquad \text{for } g \geq 2.
\end{aligned}$$

Moreover, such an allocation can be found in time $O(n^{q(g)} + gn^2 \log n)$.

We note that the constant 4 in this likely not tight. However, since it has little practical bearing on our result, both for small g and asymptotically, we do not attempt to optimize it.

Proof. We proceed by induction on g. In the $g = 1$ case there is nothing to show: we simply pick the best allocation for the single group of students and it is obviously proportional.

For $n \geq 2$, we first apply the inductive hypothesis to obtain a pair of allocations A and B, where A is $h\left(\lceil \frac{g}{2} \rceil\right)$-proportional for the first $\lceil \frac{g}{2} \rceil$ groups and B is $h\left(\lfloor \frac{g}{2} \rfloor\right)$-proportional for the remaining $\lfloor \frac{g}{2} \rfloor$ groups. Without loss of generality, we assume all schools have capacity 1; we can achieve this by turning each school of capacity c into c schools with capacity 1. For schools that are over-allocated beyond capacity, we additionally create extra "pretend" schools of capacity 1.

Let $\mathcal{H}$ be the bipartite graph with vertices for every student and every school (including the pretend schools), and edges comprised of the union of all edges that are in A or B but not both. Observe that A and B are matchings in $\mathcal{H}$. Thus, $\mathcal{H}$ has maximum degree 2, so it decomposes into a union of cycles, paths, and isolated vertices. We next perform a sequence of operations on $\mathcal{H}$ to create a new graph $\mathcal{H}'$ obtained from modified matchings A' and B'. The goal is to construct $\mathcal{H}'$ such that paths terminate in schools of value zero. Through each of these operations, we preserve the property that $u(i, A') = u(i, A)$ for all $i \leq \lceil \frac{g}{2} \rceil$ and $u(i, B') = u(i, B)$ for all $i > \lceil \frac{g}{2} \rceil$.

Suppose that a path terminates in a school k matched via A to a student j belonging to one of the first $\lceil \frac{g}{2} \rceil$ groups. This means student j is matched to a different school in B. We then change B so that both allocations match j to k. In the resulting graph, this shortens the path by two edges. Similarly, if a path terminates in a school k matched via B to a student j belonging to one of the last $\lfloor \frac{g}{2} \rfloor$ groups, we change A so that both allocations match j to k, again shortening the path by two. We iteratively continue this process until all remaining paths terminate in either a school k matched via B to a student j belonging to one of the first $\lceil \frac{g}{2} \rceil$ groups or matched via A to a student j belonging to one of the last $\lfloor \frac{g}{2} \rfloor$ groups. For each such edge, we re-match j so that it is matched to a new imaginary school of value zero instead of k (with the understanding that j can always be matched to k at the end instead of the imaginary school, for only a possible increase in utility). It is straightforward to observe that the resulting allocations A' and B' and graph $\mathcal{H}'$ have the required properties.

We relabel the schools of degree 2 in $\mathcal{H}'$ as $1, 2, 3, \ldots, m'$ in a way that consecutively lists schools from each cycle/path in $\mathcal{H}'$, one component after another. Under this labeling, for any group $i \in [g]$ and school $k \in [m']$, let

$$u_{A'}(i,k) := \begin{cases} v_k & \text{if } A' \text{ matches school } k \text{ to group } i \\ 0 & \text{otherwise} \end{cases},$$

$$u_{B'}(i,k) := \begin{cases} v_k & \text{if } B' \text{ matches school } k \text{ to group } i \\ 0 & \text{otherwise} \end{cases}.$$

We define a g-player instance of the cake-frosting problem $(f_1, f_2, \ldots, f_g, r)$ as follows. For each $i \in [g]$ we let

$$f_i(x) := u_{B'}(i, \lceil m'x \rceil) - u_{A'}(i, \lceil m'x \rceil).$$

Finally, we set

$$r := \left\lfloor \frac{g}{2} \right\rfloor.$$

Note that f is piecewise-constant with m' pieces, so axioms (1) and (2) hold. To see that axiom (3) holds, observe that

$$\sum_{i=1}^{g} f_i(x) = \sum_{i=1}^{g} u_{B'}(i, \lceil m'x \rceil) - \sum_{i=1}^{g} u_{A'}(i, \lceil m'x \rceil).$$

Note that each sum in the right-hand side has only one nonzero element, coming from the group i to which the relevant allocation matches school $\lceil m'x \rceil$. The value is $v_{\lceil m'x \rceil}$ in both sums, so they cancel out and the total sum is zero as desired.

Let X be a perfect frosting with $q(g)$ cuts. We then define a school assignment C by considering the following disjoint sets of students. Let S_0 be the set of students with degree zero in $\mathcal{H}'$, i.e., where matchings A' and B' agree. Let S_1 be the set of students neighboring two schools k_1 and k_2 such that $[\frac{k_1}{m'}, \frac{k_1+1}{m'}) \cup [\frac{k_2}{m'}, \frac{k_2+1}{m'}) \subseteq X$, and S_2 be the set of students neighboring two schools k_1 and k_2 such that $[\frac{k_1}{m'}, \frac{k_1+1}{m'}) \cup [\frac{k_2}{m'}, \frac{k_2+1}{m'}) \subseteq [0,1] \setminus X$. Finally, let $S_3 := N \setminus (S_0 \cup S_1 \cup S_2)$ be the remaining set of students. For $j \in S_0$, we assign $C(j) = A'(j) = B'(j)$, for $j \in S_1$ we assign $C(j) = A'(j)$, and for $j \in S_2$ we assign $C(j) = B'(j)$. For $j \in S_3$, we assign $C(j)$ to either school in such a way as to minimize capacity violations. See Figure 3.

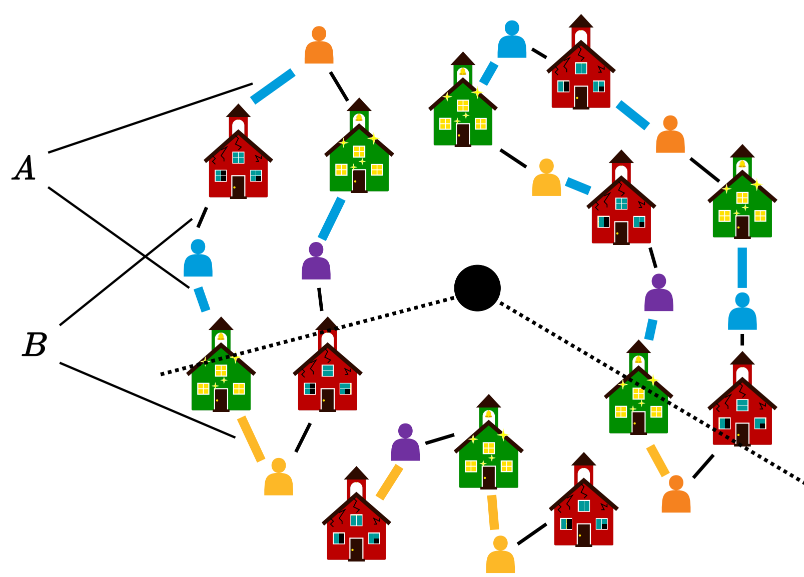


Figure 3: Illustration of how to use a perfect frosting to define an allocation for $g = 4$ groups. We suppose we have already recursively found an allocation A that is preferred by the blue and purple groups and an allocation B that is preferred by the orange and yellow groups. The edges shown compose the graph $\mathcal{H}$ arranged into a circular "cake"; the dashed line represents a division of the cake into an unfrosted part above and a frosted part below. Each cut splits a cycle at most twice and a path at most once, guaranteeing that there are at most two capacity violations per cut.

Note that there are three kinds of capacity violations that may occur in C:

- Previously existing capacity violations from A', which are also capacity violations in A, of which there are at most $h\left(\lceil \frac{g}{2} \rceil\right)$.
- Previously existing capacity violations from B', which are also capacity violations in B, of which there are at most $h\left(\lfloor \frac{g}{2} \rfloor\right)$.
- Places where there is a transition between A' and B' and some school is assigned both neighboring students, one from A' and one from B' (as is the case for both good schools intersecting the dashed line in Figure 3). Since students are enumerated in order along paths and cycles, there are only two ways this could happen.

  (1) For each cut at a school k, the students in S_3 on either side of k may both need to be assigned to k.

  (2) For each cycle that is cut at least once, there may be a similar inconsistency where the cycle wraps around.

  In the worst case, each cut occurs in a distinct cycle, yielding two capacity violations, one of each type above. We can thus bound the total number of capacity violations due to transitions between A' and B' by $2q(g)$.

In total, we have shown that

$$cv(C) \leq h\left(\left\lceil \frac{g}{2} \right\rceil\right) + h\left(\left\lfloor \frac{g}{2} \right\rfloor\right) + 2q(g).$$

Note that this is better than the required bound by $2q(g)$.

All that remains is to show that C is $h(g)$-proportional. To that end, denote by M_i an ideal allocation for group i that satisfies capacities, and let $v^\star$ be the maximum value of any school. Our objective is to show that

$$u(i, C) + h(g)v^\star \geq \frac{1}{n} u(i, M_i).$$

For any group i and $\ell \in \{0, 1, 2, 3\}$, let $u_\ell(i)$ denote the sum of utilities of group i coming from students in S_ℓ. For notational convenience, just as we partitioned the students into parts, we analogously partition the m' schools of degree 2 in $\mathcal{H}'$ into parts P_1, P_2, and P_3, where P_1 contains the schools for which both neighbors in $\mathcal{H}'$ are in S_1, P_2 contains the schools for which both neighbors in $\mathcal{H}'$ are in S_2, and P_3 contains all remaining students. For any group i,

$$\begin{aligned}
u(i, C) + h(g)v^\star &= u_0(i) + u_1(i) + u_2(i) + u_3(i) + h(g)v^\star \\
&= u_0(i) + u_3(i) + h(g)v^\star + \sum_{k \in P_1} u_{A'}(i, k) + \sum_{k \in P_2} u_{B'}(i, k) \\
&= u_0(i) + u_3(i) + h(g)v^\star + \sum_{k \in P_1} \int_{\frac{k-1}{m'}}^{\frac{k}{m'}} u_{A'}(i, k) dx + \sum_{k \in P_2} \int_{\frac{k-1}{m'}}^{\frac{k}{m'}} u_{B'}(i, k) dx \\
&= u_0(i) + u_3(i) + h(g)v^\star + m' \int_{x \in \bigcup_{k \in P_1} [\frac{k-1}{m'}, \frac{k}{m'})} u_{A'}(i, k) dx \\
&\quad + m' \int_{x \in \bigcup_{k \in P_2} [\frac{k-1}{m'}, \frac{k}{m'})} u_{B'}(i, k) dx
\end{aligned}$$

$$
\begin{aligned}
&\geq u_0(i) + (h(g) - 4q(g))v^\star + m' \int_{x\in[0,1]\setminus X} u_{A'}(i, \lceil m'x \rceil)dx + m' \int_{x\in X} u_{B'}(i, \lceil m'x \rceil)dx \\
&= u_0(i) + \left(h\left(\left\lceil \frac{g}{2} \right\rceil \right) + h\left(\left\lfloor \frac{g}{2} \right\rfloor \right) \right) v^\star + m' \int_{x\in[0,1]} u_{A'}(i, \lceil m'x \rceil)dx \\
&\qquad + m' \int_{x\in X} (u_{B'}(i, \lceil m'x \rceil) - u_{A'}(i, \lceil m'x \rceil))dx \\
&= u_0(i) + \left(h\left(\left\lceil \frac{g}{2} \right\rceil \right) + h\left(\left\lfloor \frac{g}{2} \right\rfloor \right) \right) v^\star + m' \int_{x\in[0,1]} u_{A'}(i, \lceil m'x \rceil)dx + m' \int_{x\in X} f_i(x)dx \\
&= u_0(i) + \left(h\left(\left\lceil \frac{g}{2} \right\rceil \right) + h\left(\left\lfloor \frac{g}{2} \right\rfloor \right) \right) v^\star + m' \int_{x\in[0,1]} u_{A'}(i, \lceil m'x \rceil)dx \\
&\qquad + \frac{\left\lfloor \frac{g}{2} \right\rfloor}{g} m' \int_{x\in[0,1]} f_i(x)dx \qquad \text{(because X is a perfect frosting)} \\
&= u_0(i) + \left(h\left(\left\lceil \frac{g}{2} \right\rceil \right) + h\left(\left\lfloor \frac{g}{2} \right\rfloor \right) \right) v^\star + m' \int_{x\in[0,1]} u_{A'}(i, \lceil m'x \rceil)dx \\
&\qquad + \frac{\left\lfloor \frac{g}{2} \right\rfloor}{g} m' \int_{x\in[0,1]} (u_{B'}(i, \lceil m'x \rceil) - u_{A'}(i, \lceil m'x \rceil))dx \\
&= \left(\frac{\left\lceil \frac{g}{2} \right\rceil}{g} \left(m' \int_{x\in[0,1]} u_{A'}(i, \lceil m'x \rceil)dx + u_0(i) \right) + h\left(\left\lceil \frac{g}{2} \right\rceil \right) \right) \\
&\qquad + \left(\frac{\left\lfloor \frac{g}{2} \right\rfloor}{g} \left(m' \int_{x\in[0,1]} u_{B'}(i, \lceil m'x \rceil)dx + u_0(i) \right) + h\left(\left\lfloor \frac{g}{2} \right\rfloor \right) \right) \\
&= \left(\frac{\left\lceil \frac{g}{2} \right\rceil}{g} u(i, A') + h\left(\left\lceil \frac{g}{2} \right\rceil \right) v^\star \right) + \left(\frac{\left\lfloor \frac{g}{2} \right\rfloor}{g} u(i, B') + h\left(\left\lfloor \frac{g}{2} \right\rfloor \right) v^\star \right).
\end{aligned}
$$

Note that the one inequality in the derivation above holds because there are at most 4 students in S_3 for each of the $q(g)$ cuts whose utilities are double-counted in the Cake-Frosting instance. By adding in the utility of the best school for each of these students, we compensate for this double-counting.

If i is one of the first $\left\lceil \frac{g}{2} \right\rceil$ groups, then, using the fact that $u(i, A') = u(i, A)$, we can bound the above expression by

$$
\begin{aligned}
\left(\frac{\left\lceil \frac{g}{2} \right\rceil}{g} u(i, A') + h\left(\left\lceil \frac{g}{2} \right\rceil \right) v^\star \right) + \left(\frac{\left\lfloor \frac{g}{2} \right\rfloor}{g} u(i, B') + h\left(\left\lfloor \frac{g}{2} \right\rfloor \right) v^\star \right) &\geq \left(\frac{\left\lceil \frac{g}{2} \right\rceil}{g} u(i, A) + h\left(\left\lceil \frac{g}{2} \right\rceil \right) v^\star \right) \\
&\geq \frac{\left\lceil \frac{g}{2} \right\rceil}{g} \left(\frac{1}{\left\lceil \frac{g}{2} \right\rceil} u(i, M_i) \right) \\
&\qquad \text{(by the inductive hypothesis)} \\
&= \frac{1}{g} u(i, M_i).
\end{aligned}
$$

Analogously, if i is one of the last $\left\lfloor \frac{g}{2} \right\rfloor$ groups, then we can bound the expression by

$$
\left(\frac{\left\lceil \frac{g}{2} \right\rceil}{g} u(i, A') + h\left(\left\lceil \frac{g}{2} \right\rceil \right) v^\star \right) + \left(\frac{\left\lfloor \frac{g}{2} \right\rfloor}{g} u(i, B') + h\left(\left\lfloor \frac{g}{2} \right\rfloor \right) v^\star \right) \geq \left(\frac{\left\lfloor \frac{g}{2} \right\rfloor}{g} u(i, B) + h\left(\left\lfloor \frac{g}{2} \right\rfloor \right) v^\star \right)
$$

$$\geq \frac{\lfloor \frac{g}{2} \rfloor}{g} \left(\frac{1}{\lfloor \frac{g}{2} \rfloor} u(i, M_i) \right)$$

(by the inductive hypothesis)

$$= \frac{1}{g} u(i, M_i).$$

Thus, in either case, we have shown that $u(i, C) \geq \frac{1}{g} u(i, M_i)$, concluding the existence proof.

For the runtime claim, observe that the base case simply requires finding a maximum weight matching for each group separately, which can be accomplished by the Hungarian algorithm [Kuhn, 1955, Munkres, 1957]. From our assumptions on $\mathcal{G}$, this will take time $O(n^2 \log n)$ per group. If any recursive cases take more time than this, then the bulk of the computational work is finding a perfect frosting among all g groups at the very end, after all recursive calls have finished. We can accomplish this by checking all possible placements of the at most $q(n)$ cuts, taking time $O(n^{q(n)})$. Thus, in total, the runtime is $O(n^{q(n)} + gn^2 \log n)$ □

Proof of Theorem 3. Follows immediately from Lemma 4, Lemma 5, and the master theorem. □

5 Efficient Allocations for Two or Three Groups

By Lemma 5, if we can reduce the number of cuts, we simultaneously achieve a better proportionality approximation and a faster runtime. Unfortunately, the bound in Lemma 4 of $4g - 2$ leads to a fairly unpleasant (even if polynomial) time complexity even for small g, as this becomes the degree of the polynomial. While $4g - 2$ is the best general bound we could establish for arbitrary g, it is not necessarily tight for any g. One of the reasons it may fail to be tight is that the proof of Lemma 4 reduces the cake-frosting problem to consensus halving without using the additional axioms (2) and (3) we may assume in cake-frosting. In this section we show how to use these axioms to obtain tighter bounds for $g \in \{2, 3\}$. We begin with $g = 2$, which follows from a simple application of the intermediate value theorem.

Theorem 6. *In any instance of the school redistricting problem with $g = 2$ groups, there exists a 4-proportional allocation that can be found in time $O(n^2 \log n)$.*

Proof. In light of Lemma 5, it suffices to show that we can attain $q(2) = 1$, i.e., any instance of the cake-frosting problem with two players has a perfect frosting with one cut. Consider an arbitrary instance $(f_1, f_2, 1)$, and define a function $f : [0, 1] \to \mathbb{R}$ by

$$f(x) := \int_0^x f_1(x) dx.$$

Since $f(0) = 0$, $f(1) = v_1([0, 1])$, and f is continuous, there must exist some x^* such that $f(x^*) = \frac{1}{2} v_1([0, 1])$, which means that

$$v_1([0, x^*]) = \int_0^{x^*} f_1(x^*) dx = f(x^*) = \frac{1}{2} v_1([0, 1]).$$

Furthermore, it follows from axiom (3) that

$$v_2([0, x^*]) = -v_1([0, x^*]) = -\frac{1}{2} v_1([0, 1]) = \frac{1}{2} v_2([0, 1]).$$

Thus, $[0, x^*]$ is a perfect frosting with one cut (at x^*). □

We now turn to the case of $g = 3$, which is significantly more involved. The remainder of this section is dedicated to proving the following theorem, which is the second and final main result of our paper.

Theorem 7. *In any instance of the school redistricting problem with $g = 3$ groups, there exists a 12-proportional allocation that can be found in time $O(n^2 \log n)$.*

To prove this, we require an adaptation of the Mountain Climbing Lemma [Goodman et al., 1989]. Informally, this says that two climbers can summit a two dimensional drawing of a mountain in the plane while maintaining the same y coordinate (see Figure 4).

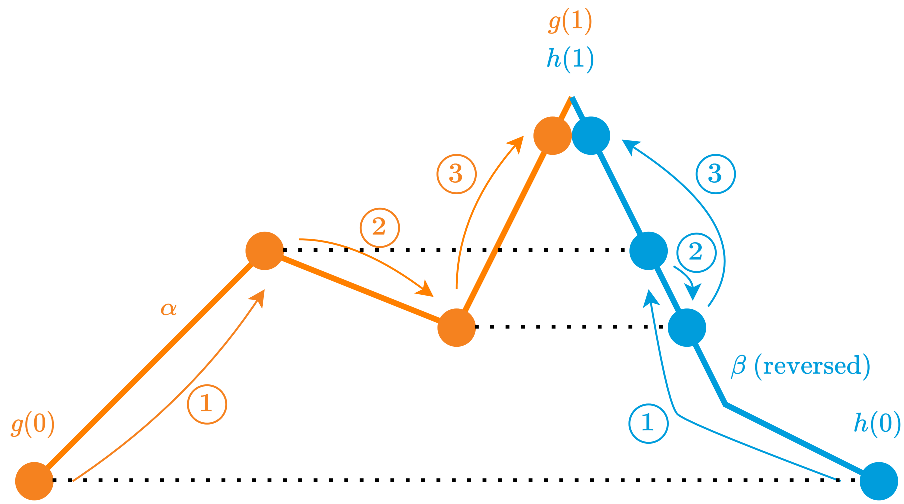


Figure 4: Depiction of the Mountain Climbing Lemma. The orange climber must summit the mountain defined by α from the left, while the blue climber must summit the mountain defined by β from the right, always maintaining the same elevation.

Lemma 8 (Mountain Climbing Lemma)**.** *Given real values $g(0) < g(1)$ and $h(0) < h(1)$, let $\alpha : [g(0), g(1)] \to \mathbb{R}$ and $\beta : [h(0), h(1)] \to \mathbb{R}$ be piecewise-linear functions such that:*

- $\alpha(g(0)) = \beta(h(0))$.
- $\alpha(g(1)) = \beta(h(1))$.
- *For all $x \in (g(0), g(1))$, $\alpha(g(0)) < \alpha(x) < \alpha(g(1))$.*
- *For all $x \in (h(0), h(1))$, $\beta(h(0)) < \beta(x) < \beta(h(1))$.*

Then it is possible to extend g and h into continuous functions over $[0, 1]$ such that, for all $t \in [0, 1]$, $\alpha(g(t)) = \beta(h(t))$.

We require a novel variant of this result. Instead of a mountain, the climbers now face an infinitely increasing ridge. The ridge is periodic in a sense, rising one vertical unit for every horizontal unit traveled, and repeating the same pattern every one horizontal unit. The following result, which we term the Ridge Climbing Lemma, says that two climbers can climb any portion of the ridge of arbitrary length while staying within one horizontal unit and maintaining the same vertical distance; see Figure 5 for an illustration.

Lemma 9 (Ridge Climbing Lemma)**.** *Let $f : \mathbb{R} \to \mathbb{R}$ be a function such that:*

- *For any $x \in \mathbb{R}$ and $n \in \mathbb{Z}$, $f(x + n) = f(x) + n$.*
- *f is continuous and piecewise linear over $[0, 1]$.*

Then, for any $a < b \in \mathbb{R}$ and $s \in (0, 1)$, there exist continuous functions $g, h : [0, 1] \to \mathbb{R}$ such that, for all $t \in [0, 1]$, the following hold:

(1) $g(t) < h(t) < g(t) + 1$

(2) $f(h(t)) - f(g(t)) = s$

(3) $g(0) < h(0) \leq a < b \leq g(1) < h(1)$

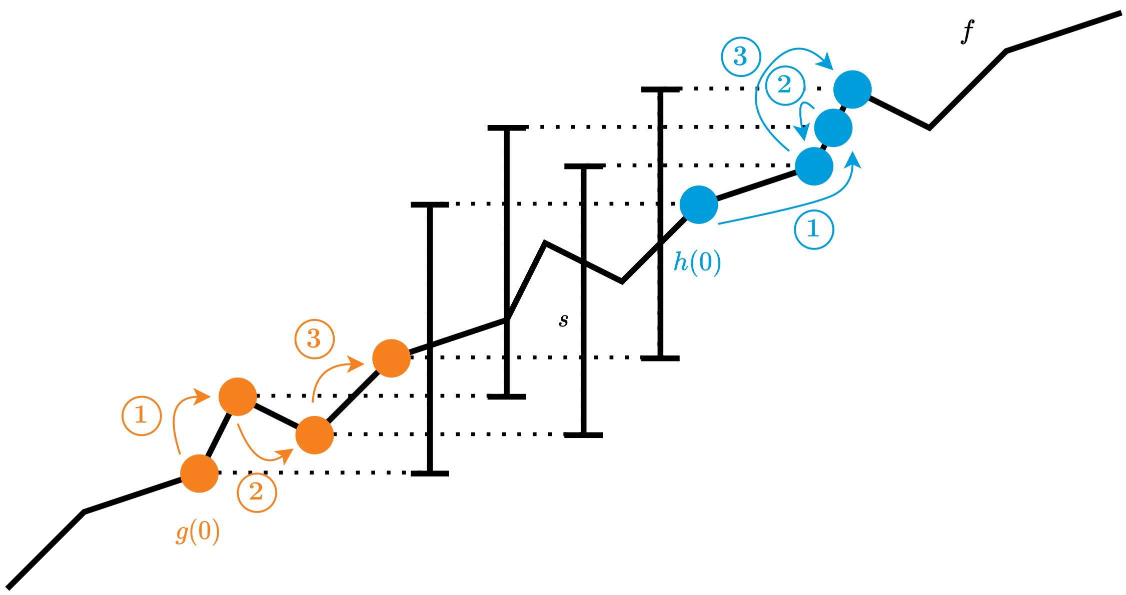


Figure 5: Depiction of the Ridge Climbing Lemma. Both climbers must climb the same periodically increasing function f while maintaining the same difference s in elevation.

Proof. Restricted to $[0, 1]$, the function f is a continuous function on a compact set, so it attains a minimum value ℓ and a maximum value u. We begin by defining g and h at the endpoints 0 and 1, then apply the Mountain Climbing Lemma to extend g and h to the entire interval. We let

$$\begin{aligned} g(0) &:= \max\{x \in \mathbb{R} \mid f(x) \leq a + \ell - 1 - s\}, \\ h(0) &:= \max\{x \in \mathbb{R} \mid f(x) \leq a + \ell - 1\}, \\ g(1) &:= \min\{x \in \mathbb{R} \mid f(x) \geq b + u\}, \end{aligned}$$

$$h(1) := \min\{x \in \mathbb{R} \mid f(x) \geq b + u + s\}.$$

We must argue that each max and min above is well-defined. From continuity, we know that, for any given $y \in \mathbb{R}$, the sets

$$S^- := \{x \in \mathbb{R} \mid f(x) \leq y\}$$
$$S^+ := \{x \in \mathbb{R} \mid f(x) \geq y\}$$

are closed. We claim that S^- is bounded above by $y - \ell + 1$. To see this, let $x \in \mathbb{R}$ such that $x > y - \ell + 1$. We may write $x = x' + n$ where n is an integer and $x' \in [0, 1)$. Then, using the special property of f, we have

$$f(x) = f(x' + n) = f(x') + n \geq \ell + n = \ell + (x - x') > \ell + x - 1 > y,$$

so $x \notin S^-$. We may similarly see that S^+ is bounded below by $y - u$. As before, let $x \in \mathbb{R}$ be such that $x < y - u$, and write $x = x' + n$ where n is an integer and $x' \in [0, 1)$. Then

$$f(x) = f(x' + n) = f(x') + n \leq u + n = u + (x - x') \leq u + x < y,$$

so $x \notin S^+$. Finally, note that S^- being bounded above implies S^+ is nonempty, and S^+ being bounded below implies S^- is nonempty. Thus, S^- is closed, nonempty, and bounded above, so it contains a maximum value. Similarly, S^+ is closed, nonempty, and bounded below, so it contains a minimum value. This shows that $g(0)$, $h(0)$, $g(1)$, and $h(1)$ are all well-defined.

Moreover, we can use the preceeding claims to show the weak inequalities in (3) hold. Specializing $y := a + \ell - 1$, the claim about S^- implies every element of the set in the definition of $h(0)$ is upper-bounded by

$$y - \ell + 1 = a + \ell - 1 - \ell + 1 = a,$$

so $h(0) \leq a$. Specializing $y := b + u$, the claim about S^+ implies every element of the set in the definition of $g(1)$ is lower-bounded by

$$y - u = b + u - u = b,$$

so $g(1) \geq b$.

Thus, the weak inequalities of (3) hold. The first and last inequalities in (3) follow from (1), which we will prove at the very end. First, though, we turn to (2).

To apply the Mountain Climbing Lemma, we imagine a mountain where, on one side, a climber must climb the function $\alpha(x) := f(x) - f(g(0))$ from $g(0)$ to $g(1)$, and on the other side, the other climber must climb the function $\beta(x) := f(x) - f(h(0))$ from $h(0)$ to $h(1)$. Note that both of these functions start out at zero since $\alpha(g(0)) = f(g(0)) - f(g(0)) = 0$ and $\beta(h(0)) = f(h(0)) - f(h(0)) = 0$. Furthermore, they summit at the same elevation of

$$\alpha(g(1)) = f(g(1)) - f(g(0)) = (b + u) - ((a + \ell - 1) - s)$$

$$= ((b + u) + s) - (a + \ell - 1) = f(h(1)) - f(h(0)) = \beta(h(1)).$$

In between these extreme endpoints, we claim that both α and β are strictly bounded between 0 and the summit value of $(b + u) - (a + \ell - 1) + s$. There are four cases to rule out and the proofs are analogous, but we include them all nonetheless.

First, suppose toward a contradiction that $\alpha(x) \leq 0$ for some $x \in (g(0), g(1))$. This means that $f(x) \leq f(g(0)) = (a+\ell-1)-s$. However, we know that $f(g(1)) = (b+u) > (a+\ell-1) > (a+\ell-1)-s$. Therefore, by the intermediate value theorem, we know that there is some $x' \in [x, g(1)]$ such that $f(x') = (a+\ell-1)-s$. Since $x' \geq x > g(0)$, this contradicts the definition of $g(0)$. Therefore, $\alpha(x) > 0$ for all $x \in (g(0), g(1))$.

Next, suppose toward a contradiction that $\alpha(x) \geq (b+u)-(a+\ell-1)+s$ for some $x \in (g(0), g(1))$. This means that $f(x)-f(g(0)) \geq (b+u)-(a+\ell-1)+s$. Since $f(g(0)) = (a+\ell-1)-s$, this is the same as $f(x) \geq (b+u)$. However, we know that $f(g(0)) = (a+\ell-1)-s < (a+\ell-1) < (b+u)$. Therefore, by the intermediate value theorem, we know that there is some $x' \in [g(0), x]$ such that $f(x') = (b+u)$. Since $x' \leq x < g(1)$, this contradicts the definition of $g(1)$. Therefore, $\alpha(x) < (b+u)-(a+\ell-1)+s$ for all $x \in (g(0), g(1))$.

Next, suppose toward a contradiction that $\beta(x) \leq 0$ for some $x \in (h(0), h(1))$. This means that $f(x) \leq f(h(0)) = (a+\ell-1)$. However, we know that $f(h(1)) = (b+u)+s > (b+u) > (a+\ell-1)$. Therefore, by the intermediate value theorem, we know that there is some $x' \in [x, h(1)]$ such that $f(x') = (a+\ell-1)$. Since $x' \geq x > h(0)$, this contradicts the definition of $h(0)$. Therefore, $\beta(x) > 0$ for all $x \in (h(0), h(1))$.

Finally, suppose toward a contradiction that $\beta(x) \geq (b+u)-(a+\ell-1)+s$ for some $x \in (h(0), h(1))$. This means that $f(x)-f(h(0)) \geq (b+u)-(a+\ell-1)+s$. Since $f(h(0)) = (a+\ell-1)$, this is the same as $f(x) \geq (b+u)+s$. However, we know that $f(h(0)) = (a+\ell-1) < (b+u) < (b+u)+s$. Therefore, by the intermediate value theorem, we know that there is some $x' \in [h(0), x]$ such that $f(x') = (b+u)+s$. Since $x' \leq x < h(1)$, this contradicts the definition of $h(1)$. Therefore, $\beta(x) < (b+u)-(a+\ell-1)+s$ for all $x \in (h(0), h(1))$.

Thus, we may apply the Mountain Climbing Lemma to extend g and h into continuous functions over $[0,1]$ such that, for all $t \in [0,1]$, $\alpha(g(t)) = \beta(h(t))$. By the definitions of α and β, this means

$$f(g(t)) - f(g(0)) = f(h(t)) - f(h(0)),$$

or, equivalently,

$$f(h(t)) - f(g(t)) = f(h(0)) - f(g(0)) = (a+\ell-1) - ((a+\ell-1)-s) = s.$$

This establishes (2).

All that remains is to prove (1). Consider the function $z(t) := h(t) - g(t)$. We must show that, for all $t \in [0,1]$, $z(t) \in (0,1)$. Let us first consider the special case of $t = 0$. By definition,

$$z(0) = h(0) - g(0) = \max\{x \in \mathbb{R} \mid f(x) \leq (a+\ell-1)\} - \max\{x \in \mathbb{R} \mid f(x) \leq (a+\ell-1)-s\} \geq 0$$

since the second set is contained in the first set. We next claim that $z(0) \leq 1$. Suppose toward a contradiction that $z(0) = 1 + x'$ where $x' > 1$. Observe that $f(g(0)+x') > (a+\ell-1)-s$, for otherwise $g(0)+x'$ would be a larger value of x in the set

$$\{x \in \mathbb{R} \mid f(x) \leq (a+\ell-1)-s\}$$

than $g(0)$, contradicting the definition of $g(0)$ as the maximum element of that set. Therefore, using the special property of f again, we have

$$\begin{aligned}(a+\ell-1) = f(h(0)) = f(g(0)+z(0)) = f(g(0)+x'+1)\\ = f(g(0)+x')+1 > (a+\ell-1)-s+1 > (a+\ell-1),\end{aligned}$$

which is a contradiction. Thus, we have shown that $0 \le z(0) \le 1$. We next claim that z can never take the value of 0 or 1. Suppose toward a contradiction that $z(t) = 0$ for some $t \in [0,1]$. This means $h(t) = g(t)$, and thus $f(h(t)) = f(g(t))$, contradicting (2) since $s > 0$. Now suppose toward a contradiction that $z(t) = 1$ for some $t \in [0,1]$. In this case, $h(t) = g(t) + 1$, so

$$f(h(t)) = f(g(t) + 1) = f(g(t)) + 1,$$

which contradicts (2) since $s < 1$. Thus, z is a continuous function starting between 0 and 1, but never taking the value 0 or 1, so it must stay strictly between 0 and 1 for all t. This concludes the proof of (1). ☐

With the lemma in hand, we are now ready to prove Theorem 7.

Proof of Theorem 7. In light of Lemma 5, it suffices to show that we can attain $q(3) = 2$, i.e., any instance of the cake-frosting problem with three players has a perfect frosting with two cuts. (The extra cut comes from one of the two recursive cases, where we have already established via Theorem 6 that finding a perfect frosting amongst two groups can be obtained with one additional cut.)

Consider an instance $(f_1, f_2, f_3, 1)$ (the case where $r = 2$ is analogous, as it is essentially the same problem). Let $f : [0,1] \to \mathbb{R}$ be defined by

$$f(x) := \frac{v_3([0,x])}{v_3([0,1])}.$$

Observe that f is piecewise linear (since f_3 is piecewise constant by axiom (2)), $f(0) = 0$ and $f(1) = 1$. We can thus extend f to be a continuous function on all of $\mathbb{R}$ using the rule that $f(x+n) = f(x) + n$ for $n \in \mathbb{Z}$, so that f satisfies all the conditions of the Ridge Climbing Lemma.

Let $a \in [0,1]$ be any value such that there is a finite set X_0 of values x such that $f(x) = f(a)$. Such a point is guaranteed to exist since f is piecewise-linear on $[0,1]$ and not constant over the entire interval; we just need to avoid the finitely-many y-values where f is constant. Let $\ell := |X_0|$ and set $b := a + 3\ell$. Finally, set $s := \frac{1}{3}$ and apply the Ridge Climbing Lemma to obtain functions g and h.

Let $x_0 := a$, and inductively define, for each $i \ge 0$, x_{i+1} to be the value of $h(t_i)$ at the smallest t such that $g(t_i) = x_i$. From conditions (1) and (3) of the Ridge Climbing Lemma and the way we chose b, this sequence is monotonically increasing and can continue at least until $x_{3\ell}$. Observe that, for each i,

$$\begin{aligned} f(x_{i+3}) - f(x_i) &= f(x_{i+3}) - f(x_{i+2}) + f(x_{i+2}) - f(x_{i+1}) + f(x_{i+1}) - f(x_i) \\ &= f(h(t_{i+2})) - f(g(t_{i+2})) + f(h(t_{i+1})) - f(g(t_{i+1})) + f(h(t_i)) - f(g(t_i)) \\ &= \frac{1}{3} + \frac{1}{3} + \frac{1}{3} = 1. \end{aligned}$$

It inductively follows that, for any positive integer i, $f(x_{3i}) - i = f(x_0) = f(a)$. Let X_{3i} be the set of values x such that $f(x) = f(x_{3i})$. Since $f(x) - i = f(x - i)$, this is the same as the set of values x such that $f(x - i) = f(x_{3i}) - i = f(a)$. Hence, X_{3i} is just X_0 shifted to the right by i. Recalling that X_0 has ℓ elements and using the pigeonhole principle, it follows that two values from the subsequence $x_0, x_3, x_6, x_9, \ldots, x_{3\ell}$ must differ by an integer. Thus, let $i^* < j^*$ and $k \ge 1$ be such that $x_{j^*} - x_{i^*} = k$.

Note that

$$\begin{aligned}\frac{j^*-i^*}{3} &= \sum_{i=i^*}^{j^*-1} \frac{1}{3}\\ &= \sum_{i=i^*}^{j^*-1} (f(x_{i+1}) - f(x_i))\\ &= f(x_{j^*}) - f(x_{i^*})\\ &= f(x_{i^*}+k) - f(x_{i^*})\\ &= f(x_{i^*}) + k - f(x_{i^*})\\ &= k.\end{aligned}$$

Let us extend f_1 (and thus v_1) to be defined on all of $\mathbb{R}$ so that it is periodic with period 1. Since the collection $\{[x_i, x_{i+1}] \mid i^* \le i < j^*\}$ partitions an interval of length k,

$$\sum_{i=i^*}^{j^*-1} v_1([x_i, x_{i+1}]) = k \cdot v_1([0,1]) = \frac{j^*-i^*}{3} \cdot v_1([0,1]),$$

or, equivalently,

$$\frac{1}{j^*-i^*} \sum_{i=i^*}^{j^*-1} v_1([x_i, x_{i+1}]) = \frac{1}{3} \cdot v_1([0,1]).$$

By the averaging principle, one of the $j^* - i^*$ terms in the sum must be at least $\frac{1}{3}v_1([0,1])$ and another must be at most $\frac{1}{3}v_1([0,1])$. Call the first of these two indices of the sum i_1, and the second index i_2. Then consider the function $\gamma : [t_{i_1}, t_{i_2}] \to \mathbb{R}$ defined by

$$\gamma(t) := v_1([g(t), h(t)]).$$

Note that

$$\gamma(t_{i_1}) = v_1([g(t_{i_1}), h(t_{i_1})]) = v_1([x_{i_1}, x_{i_1+1}])$$

and

$$\gamma(t_{i_2}) = v_1([g(t_{i_2}), h(t_{i_2})]) = v_1([x_{i_2}, x_{i_2+1}]).$$

Since one of these is at least $\frac{1}{3}v_1([0,1])$ and the other at most $\frac{1}{3}v_1([0,1])$, the intermediate value theorem implies there is some t^* such that

$$\gamma(t^*) = \frac{1}{3}v_1([0,1]).$$

We claim that the interval $[g(t^*), h(t^*)]$ is a perfect frosting.[4] For player 1, observe that

$$v_1([g(t^*), h(t^*)]) = \gamma(t^*) = \frac{1}{3}v_1([0,1]).$$

[4] Again, this is under the interpretation that the interval $[0,1]$ "wraps around"; e.g., if $[g(t^*), h(t^*)] = [5.8, 6.1]$ then we really mean the set $[0, 0.1] \cup [0.8, 1]$. Under this interpretation of an interval, there are still only at most two cuts.

For player 3,

$$\begin{aligned}
v_3([g(t^*), h(t^*)]) &= v_3([0, h(t^*)]) - v_3([0, g(t^*)]) \\
&= f(h(t^*))v_3([0,1]) - f(g(t^*))v_3([0,1]) \qquad \text{(by the definition of } f) \\
&= (f(h(t^*)) - f(g(t^*)))v_3([0,1]) \\
&= \frac{1}{3}v_1([0,1]),
\end{aligned}$$

where in the final equality we have used condition (2) of the Ridge Climbing Lemma. Finally, for player 2, it follows from axiom (3) that

$$v_2([g(t^*), h(t^*)]) = -v_1([g(t^*), h(t^*)]) - v_3([g(t^*), h(t^*)]) = -\frac{1}{3}(v_1([0,1]) + v_3([0,1])) = \frac{1}{3}v_2([0,1]).$$

Thus, $[g(t^*), h(t^*)]$ is a perfect frosting. □

6 Open Questions and Limitations

There are several technical questions that our work leaves open. First is the existential question about our relaxed fairness notion, t-proportionality, as the number of groups grows. Theorem 3 shows that $O(g \log g)$-proportional allocations always exist, and Theorem 2 establishes a lower bound of $\Omega(g)$. While the gap is not large, it would be interesting to close it. Second, Theorem 3 only provides a polynomial-time algorithm for constant g, and the improved time complexity of $O(n^2 \log n)$ given by Theorems 6 and 7 — using an additional tool, namely the Ridge Climbing Lemma — only holds for $g = 2, 3$. Can these tools yield an algorithm that guarantees $O(g \log g)$-proportionality in time polynomial in g?

In terms of limitations, we acknowledge that proportionality may not be sufficient, in and of itself, to achieve intuitively fair outcomes in every situation. The same has been observed for envy-freeness in rent division, for example. The rent-division solution most commonly used in practice maximizes the minimum utility of any player subject to envy-freeness [Gal et al., 2017]. One could adopt a similar solution for school redistricting, with the optimization being handled by an integer programming solver.

In addition, as mentioned in Section 1, our approach does not capture all aspects of school redistricting in the real world; there are two gaps that are perhaps the most significant. First, reassigning students from one school to another can be contentious. Accounting for the previous assignment of students would amount to more complicated, heterogeneous utility functions that are not captured by our model. Second, school attendance zones are typically contiguous geographic regions. While our feasibility constraints are maximally general and can capture any geographic constraints (such as those based on distances), the subset of students assigned to a particular school in our model may not form a contiguous region. Despite these limitations, we feel our approach and results are auspicious first steps towards an understanding of provable fairness in school redistricting.

References

A. Abdulkadiroğlu, P. Pathak, and A. E. Roth. The New York City high school match. *American Economic Review*, 95(2):364–367, 2005a.

A. Abdulkadiroğlu, P. Pathak, A. E. Roth, and T. Sönmez. The Boston high school match. *American Economic Review*, 95(2):368–371, 2005b.

P. K. Agarwal, S.-H. Ko, K. Munagala, and E. Taylor. Locally fair partitioning. In *Proceedings of the 36th AAAI Conference on Artificial Intelligence (AAAI)*, pages 4752–4759, 2022.

M. Allman, I. Ashlagi, I. Lo, J. Love, K. Mentzer, L. Ruiz-Setz, and H. O'Connell. Designing school choice for diversity in the San Francisco Unified School District. In *Proceedings of the 23rd ACM Conference on Economics and Computation (EC)*, pages 290–291, 2022.

N. Benabbou, M. Chakraborty, E. Elkind, and Y. Zick. Fairness towards groups of agents in the allocation of indivisible items. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 95–101, 2019.

N. Benabbou, M. Chakraborty, A. Igarashi, and Y. Zick. Finding fair and efficient allocations when valuations don't add up. In *Proceedings of the 13th International Symposium on Algorithmic Game Theory (SAGT)*, pages 32–46, 2020.

G. Benadè, A. D. Procaccia, and J. Tucker-Foltz. You can have your cake and redistrict it too. In *Proceedings of the 24th ACM Conference on Economics and Computation (EC)*, 2023. Forthcoming.

S. E. Black. Do better schools matter? Parental valuation of elementary education. *Quarterly Journal of Economics*, 114(2):577–599, 1999.

S. J. Brams. Making partisan gerrymandering fair: One old and two new methods. *Social Science Quarterly*, 101(1):68–72, 2020.

E. Budish. The combinatorial assignment problem: Approximate competitive equilibrium from equal incomes. *Journal of Political Economy*, 119(6):1061–1103, 2011.

I. Caragiannis, D. Kurokawa, H. Moulin, A. D. Procaccia, N. Shah, and J. Wang. The unreasonable fairness of maximum Nash welfare. *ACM Transactions on Economics and Computation*, 7(3): article 12, 2019.

F. Caro, T. Shirabe, M. Guignard, and A. Weintraub. School redistricting: Embedding GIS tools with integer programming. *Journal of the Operational Research Society*, 55(8):836–849, 2004.

S. Clarke and S. Jurkis. An operations research approach to racial desegregation of school systems. *Socio-Economic Planning Sciences*, 1(3):259–272, 1968.

V. Conitzer, R. Freeman, and N. Shah. Fair public decision making. In *Proceedings of the 18th ACM Conference on Economics and Computation (EC)*, pages 629–646, 2017.

S. A. Cordes, C. Rick, and A. E. Schwartz. Do long bus rides drive down academic outcomes? *Educational Evaluation and Policy Analysis*, 44(4):689–716, 2022.

J. De Silva, B. Gales, B. Kagy, and D. Offner. An analysis of a fair division protocol for drawing legislative districts. arXiv:1811.05705, 2018.

A. D. Franklin and E. Koenigsberg. Computed school assignments in a large district. *Operations Research*, 21(2):413–426, 1973.

Y. Gal, M. Mash, A. D. Procaccia, and Y. Zick. Which is the fairest (rent division) of them all? *Journal of the ACM*, 64(6): article 39, 2017.

N. Gillani, D. Beeferman, C. Vega-Pourheydarian, C. Overney, P. Van Hentenryck, and D. Roy. Redrawing attendance boundaries to promote racial and ethnic diversity in elementary schools. arXiv:2303.07603, 2023.

J. E. Goodman, J. Pach, and C. K. Yap. Mountain climbing, ladder moving, and the ring-width of a polygon. *American Mathematical Monthly*, 96(6):494–510, 1989.

S.-H. Ko, E. Taylor, P. K. Agarwal, and K. Munagala. All politics is local: Redistricting via local fairness. In *Proceedings of the 35th Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2022.

H. W. Kuhn. The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2(1-2):83–97, 1955.

Z. Landau and F. E. Su. Fair division and redistricting. In *The Mathematics of Decisions, Elections, and Games*, pages 17–36. American Mathematical Society, 2014.

Z. Landau, O. Reid, and I. Yershov. A fair division solution to the problem of redistricting. *Social Choice and Welfare*, 32(3):479–492, 2009.

R. S. Liggett. The application of an implicit enumeration algorithm to the school desegregation problem. *Management Science*, 20(2):159–168, 1973.

J. Munkres. Algorithms for the assignment and transportation problems. *Journal of the Society for Industrial and Applied Mathematics*, 5(1):32–38, 1957.

W. Pegden, A. D. Procaccia, and D. Yu. A partisan districting protocol with provably nonpartisan outcomes. arXiv:1710.08781, 2017.

A. D. Procaccia. Cake cutting algorithms. In F. Brandt, V. Conitzer, U. Endress, J. Lang, and A. D. Procaccia, editors, *Handbook of Computational Social Choice*, chapter 13. Cambridge University Press, 2016.

T. Roughgarden. Resource augmentation. In T. Roughgarden, editor, *Beyond the Worst-Case Analysis of Algorithms*, chapter 4. Cambridge University Press, 2020.

F. W. Simmons and F. E. Su. Consensus-halving via theorems of Borsuk-Ulam and Tucker. *Mathematical Social Sciences*, 45(1):15–25, 2003.

W. Stromquist and D. R. Woodall. Sets on which several measures agree. *Journal of Mathematical Analysis and Applications*, 108(1):241–248, 1985.

W. Suksompong. Constraints in fair division. *SIGecom Exchanges*, 19(2):46–61, 2021.